



People`s Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
University of Echahid Hamma Lakhdar El Oued



Faculty of Technology
Department of Electrical Engineering

Dissertation

ACADEMIC MASTER

Domain: Science and Technology

Division: Electrical Engineering

Specialty: Telecommunication Systems

Presented by:

- 1.LAKHDARI Nesrine
- 2.MOUSSAOUI Raihana

Entitled:

**Image Classification Using Transfer Learning and
Training Hyperparameter Optimization**

Publicly defended in: 22 /05 /2026

Board of Examiners:

Dr. Ghandir El Saïd

Chairman

Dr.Tidjani Amina

Supervisor

Dr. Blelili Noor Al-Iman

Examiner

Academic Year :2025/2026



Abstract

Brain tumor classification from Magnetic Resonance Imaging (MRI) is an important research area in medical image analysis. In this study, several deep learningbased Transfer Learning (TL) models were investigated for brain tumor classification using MRI images. Pretrained Convolutional Neural Network (CNN) architectures, including ResNet18, ResNet50, MobileNetV2, and DenseNet201, were evaluated and compared in terms of classification performance.

To enhance the training process and improve model generalization, image preprocessing and data augmentation techniques were applied. In addition, different training hyperparameters, including learning rate, batch size, optimizer configuration, and number of epochs, were adjusted during the experimentation phase.

The performance of the evaluated models was assessed using several metrics, including accuracy, recall, F1score, and confusion matrix analysis. Experimental results showed that DenseNet201 achieved the highest classification performance, obtaining an accuracy exceeding 99.30% on the MRI dataset.

The obtained results demonstrate the effectiveness of Transfer Learning and pretrained CNN architectures for MRI brain tumor classification and highlight the impact of training parameter selection on classification performance.

Keywords:

Artificial Intelligence, Deep Learning, Convolutional Neural Networks, Transfer Learning, Magnetic Resonance Imaging, Brain Tumor Classification.

Résumé

La classification des tumeurs cérébrales à partir d'images d'Imagerie par Résonance Magnétique (IRM) constitue un domaine de recherche important dans l'analyse d'images médicales. Dans cette étude, plusieurs modèles d'apprentissage profond basés sur l'Apprentissage par Transfert (TL) ont été étudiés pour la classification des tumeurs cérébrales à partir d'images IRM. Des architectures de Réseaux de Neurones Convolutifs (CNN) pré-entraînées, notamment ResNet18, ResNet50, MobileNetV2 et DenseNet201, ont été évaluées et comparées en termes de performances de classification.

Afin d'améliorer le processus d'entraînement et de renforcer la capacité de généralisation des modèles, des techniques de prétraitement des images et d'augmentation des données ont été appliquées. De plus, différents hyperparamètres d'entraînement, notamment le taux d'apprentissage, la taille des lots, la configuration de l'optimiseur et le nombre d'époques, ont été ajustés lors de la phase d'expérimentation.

Les performances des modèles évalués ont été mesurées à l'aide de plusieurs métriques, incluant la précision, le rappel, le score F1 et l'analyse de la matrice de confusion. Les résultats expérimentaux ont montré que DenseNet201 a atteint les meilleures performances de classification, avec une précision dépassant 99.30 % sur le jeu de données IRM.

Les résultats obtenus démontrent l'efficacité de l'Apprentissage par Transfert et des architectures CNN pré-entraînées pour la classification des tumeurs cérébrales par IRM, et soulignent l'impact du choix des paramètres d'entraînement sur les performances de classification.

mot de cle :

Artificial Intelligence, Deep Learning, Convolutional Neural Networks, Transfer Learning, Magnetic Resonance Imaging, Brain Tumor Classification.

الملخص

يُعدّ تصنيف أورام الدماغ من صور الرنين المغناطيسي (MRI) من المجالات البحثية الهامة في تحليل الصور الطبية. في هذه الدراسة، تم التحقيق في عدة نماذج للتعلّم العميق القائم على نقل التعلّم (TL) لتصنيف أورام الدماغ باستخدام صور الرنين المغناطيسي. وقد تم تقييم ومقارنة بنيات الشبكات العصبية التلافيفية (CNN) المُدرّبة مسبقاً، والتي تشمل: ResNet18 و ResNet50 و MobileNetV2 و DenseNet201، من حيث أداء التصنيف.

لتعزيز عملية التدريب وتحسين قدرة النماذج على التعميم، تم تطبيق تقنيات المعالجة المسبقة للصور وتوسيع البيانات. علاوةً على ذلك، تم ضبط معاملات التدريب المختلفة خلال مرحلة التجريب، بما فيها: معدل التعلّم، وحجم الدفعة، وإعداد المُحسّن، وعدد جُوب التدريب.

تم تقييم أداء النماذج المُختبّرة باستخدام عدة مقاييس، تشمل: الدقة، والاستدعاء، ومقياس F1، وتحليل مصفوفة الالتباس. وأظهرت النتائج التجريبية أن نموذج DenseNet201 حقّق أعلى أداء في التصنيف، إذ تجاوزت دقته 99.30% على مجموعة بيانات الرنين المغناطيسي.

تُثبت النتائج المُحققة فاعلية نقل التعلّم وبنيات الشبكات العصبية التلافيفية المُدرّبة مسبقاً في تصنيف أورام الدماغ من صور الرنين المغناطيسي، كما تُبرز الأثر الكبير لاختيار معاملات التدريب على أداء التصنيف.

الكلمات المفتاحية :

الذكاء الاصطناعي التعلّم العميق الشبكات العصبية الالتفافية، التعلّم بالنقل، التصوير بالرنين المغناطيسي، تصنيف أورام الدماغ.

إهداء

(وَأَخِرُ دَعْوَاهُمْ أَنِ الْحَمْدُ لِلَّهِ رَبِّ الْعَالَمِينَ)

الحمد لله حبا وشكرا وامتنانا ماكنت لأفعل هذا لو لا فضل الله فالحمد لله على البدء وعلى الختام

أهدي هذا النجاح الى نفسي الطموحة جدا لقد ظننت انني لا أستطيع ولكن من قال انا لها نالها وان ابت رغما عنها اتيت بها وها انا اليوم اختتم بحث تخرجي بكل همة ونشاط فالحمد لله اللهم لا تجعله اخر عهدي من العلم واجعلها خير بداية لطريق أعظم اللهم بارك لنا في عملنا وانفعنا بما علمتنا.

اهدي ثواب هذا البحث إلى من تربيته على يده ومن علمني القيم والمبادئ الى من لا ينفصل اسمه عن اسمي الي فرحتي الدائمة الى

مصدر قوتي فخري لطالما عاهدته بهذا النجاح وها انا اتممت وعدي وأهديته اليك

(ابي الغالي حفظه الله)

واهدي ثمرة جهدي ودراستي وفرحتي المنتظرة الى نبراس ايامي ووهج حياتي الى التي ظلت دعواتها تضم اسمي دائما الى من

افنت عمرها في سبيل ان احقق طموحي واحلق في أعالي المراتب

(امي الغالية حفظها الله)

والى ملهمي نجاحي صناع قوتي صفوة ايامي سلوة اوقاتي الى الشموع التي تنير لي الطريق دوما

(أخي وأخواتي)

إلى روح رحلت قبل أن ترى حلمي يتحقق، وإلى من كان سنداً لي وشجعني على إكمال هذه المرحلة.

ها أنا اليوم أكمل دربي كما تمنى لي وهذه خطوتي الأولى وفاءً لوعدك .

(خالي الغالي رحمه الله)

ما سلكنا البدايات الا بتيسيره وما بلغنا النهايات الا بتوقيقه وما حققنا الغايات الا بفضله فالحمد لله

الأخضاري نسرين

إهداء

إلى روح النبي الكريم محمد صلى الله عليه وسلم

سيد الخلق، ومعلّم الإنسانية، ومَن أثار الدرب للطالبيين

إلى الوالدين الكريمين

مَن سقياني من ينبوع الحنان، وأضاء دربي بالدعاء والتضحية والعطاء،

فكان جهدي ثمرةً لبذلهم، وكان نجاحي تاجاً على رأسيهما

إلى الإخوة والأخوات الأعزاء

رفقاء العمر وسكّان القلب، الذين آزروني وشدّوا من عزيمتي في كل خطوة

إلى الصديقات الوفيّات

اللواتي كنّ سنداً وسماءً صافيةً، تمنحنني الدفء والتشجيع في أوقات الشدة والضيق

إلى الأعمام والأخوال الكرام

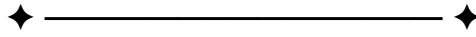
أولياء النعمة وقدوة الأجيال، الذين غرسوا فيّ قيم العلم والمثابرة والأصالة

إلى أساتذتي الأجلاء

الذين أضاءوا عقلي بنور المعرفة، ومنحوني من علمهم وصبرهم وحسن توجيههم ما لا يُحصى

إلى الأمة الإسلامية جمعاء

أمة الإسلام الشامخة، أهدي هذا الجهد المتواضع راجيةً أن يكون لبننةً في صرح العلم النافع



(رَبِّ أَوْزِعْنِي أَنْ أَشْكُرَ نِعْمَتَكَ الَّتِي أَنْعَمْتَ عَلَيَّ)

موساوي ربحانة

شكر وتقدير

نتقدم بخالص الشكر والتقدير إلى الدكتورة الفاضلة تجاني أمينة،

التي كانت لنا خير موجّه وداعم طوال فترة إنجاز هذا العمل، فكل كلمات الامتنان

لا توافيها حقها على ما بذلته من جهد وعطاء، إذ كانت نورًا يضيء طريقنا،

وسندًا حقيقيًا في رحلتنا نحو النجاح،

لقد قدمت لنا نصائح قيّمة وإرشادات مفيدة، وشجعتنا باستمرار بروحها الطيبة وصبرها الكبير،

مما كان له بالغ الأثر في إتمام هذا المشروع على أفضل وجه.

نحن نشمّن كثيرًا وقتها الثمين، وتفانيها، وحسن إشرافها، ونعتز جدًا بأننا حظينا

بفرصة العمل تحت إشرافها الكريم

جعلك الله من الذين قال عنهم

{ وَرَفَعْنَاهُ مَكَانًا عَلِيًّا }

Materiel table

Introduction general.....	1
Chapter 01:	3
Problem Formulation.....	3
1.1Introduction	4
1.2Introduction to image classification:	4
1.2.1Definition and importance of image classification :.....	4
1.3 Architecture of image classification	4
1.3.1Datasets:	5
1.3.2 Preprocessing:.....	6
1.3.3 Feature extraction :	6
1.3.4 Classification and Decision :	6
1.3.5 Evaluation:.....	6
1.4 Problem Statement and Research Objectives	8
□ definition of the data scarcity :	8
1.5 Conclusion:.....	8
Chapter 02 :	9
Theoretical Background and Fundamental Concepts of AI and Learning Methods	9
2.1 Introduction	10
2.2Definition of AI :	10
2.3 Machine learning	10
2.3.1definition.....	11
2.3.2Types of machine learning	11
2.3.3 Examples and use cases	14
2.3.4How does machine learning ?work	14
2.4 Deep Learning	15
2.4.1 definition.....	15
2.4.2Types of Deep Learning	15
2.5 Transfer Learning:	21
2.5.1 definition of transfer learning	21
2.5.2 type of transfer learning.....	21
2.6 Why Transfer Learning in Medical Image Classification:	23
2.7 State of th eart.....	23
2.8 Conclusion.....	23

Chapter 03:	24
Training protocol and hyperparametre optimization	24
3.1 Introduction	25
3.2 Proposed Methodology.....	26
3.2.1 System Architecture	26
3.2.2 Dataset Description.....	27
3.2.3 Image Preprocessing and Data Augmentation.....	28
Ø Image Resizing:	28
Ø Data Augmentation :.....	28
3.2.4 Transfer Learning Models	29
3.3 Training and validation.....	31
3.3.1 Training option	31
3.3.2 Hyperparameter Optimization	31
3.3.3 Experimental results and discussion :.....	33
3.4 Impact of Hyperparameter Optimization.....	35
3.5 Conclusion.....	39
Conclusion general	40
Annex	41
MRI (Magnetic Resonance Imaging):	41
References	43

List of figures :

Figure 1.1 General architecture of image classification.	5
Figure1. 2 An illustrative diagram of Machine Learning	10
Figure2.2: Schematic representation of data flow from labeled inputs to predictions	12
Figure 2.3: Dimensionality reduction technique as part of Unsupervised Machine Learning.	12
Figure2.4: Schematic diagram of the Semisupervised learning process	13
Figure 2.5 The architectural diagram of the Reinforcement Learning system.	14
Figure2.6 CNN Architecture Diagram	18
Figure 2.7:Neural Network with Skip Connections / Residual Connections	19
Figure 2.8 Multilayer Feedforward Neural Network Diagram.....	20
Figure 2.9: Generative Adversarial Networks GAN Architecture	20
Figure 2.10 Transfer Learning Process Diagram.....	21
Figure2.11 Domain Adaptation Process.....	22
Figure 2.12:Discover patterns without using labels.	22
Figure 3.1 Workflow of the proposed brain tumor classification system.....	26
Figure3.2: Sample MRI Images from Brain Tumor Dataset.....	28
Figure 3.3 :Examples of MRI Data Augmentation	29
Figure 3.4Transfer Learning Procedure Applied to Each CNN Architecture	30
Figure 3.5 Training options	32
Figure 3.6 Training and validation accuracy and loss curves.....	33
Figure 3.7 Confusion Matrix for Brain Tumor Classification	34
Figure 3.8 Training and validation progress curves for the DenseNet201 model	37
Figure3.9: Confusion matrix of DenseNet201 for brain tumor classification.	38

List of tables:

Table 1: Performance comparison of different deep learning models.....	32
Table 2: Comparison of different hyperparameter configurations for DenseNet201	35

Introduction general

The rapid advancement of Artificial Intelligence, particularly Deep Learning techniques, has led to transformative impacts across many fields, especially in the medical domain. Automated medical image analysis has become one of the most important modern applications aimed at assisting physicians in diagnosis, reducing the time and effort required to interpret images, and improving accuracy and reliability. Among medical imaging modalities, Magnetic Resonance Imaging (MRI) stands out for its ability to provide highly detailed visualization of brain tissues, making it an essential tool for detecting and diagnosing brain tumors.

However, the manual analysis of MRI images remains a complex task that requires significant medical expertise due to the visual similarities between different tumor types and the difficulty of distinguishing them. In addition, the large number of images that must be examined increases the burden on specialists. This challenge has highlighted the importance of image classification systems based on Convolutional Neural Networks, which have demonstrated remarkable capability in automatically extracting fine features and recognizing complex patterns in images with high precision.

In this context, Convolutional Neural Networks (CNNs) form the cornerstone of computer vision applications, as they enable models to learn features directly from raw data without the need for manual feature extraction. Furthermore, the adoption of Transfer Learning allows the use of deep models pretrained on large datasets, helping to overcome the issue of limited labeled medical data, reducing training time, and enhancing model performance when applied to MRI brain images.

This dissertation aims to investigate the use of deep learning and Transfer Learning techniques for brain MRI image classification using several pretrained Convolutional Neural Network (CNN) architectures, namely ResNet18, ResNet50, MobileNetV2, and DenseNet201. The study focuses on evaluating and comparing the classification performance of these models using different evaluation metrics, including accuracy, precision, recall, F1score, and confusion matrix analysis.

In addition to the application of Transfer Learning, this work also investigates the effect of training hyperparameter optimization on the performance of the selected model. Several training parameters, including minibatch size, number of epochs, and learning rate, were adjusted in order to analyze their impact on classification accuracy and model generalization while maintaining a balance between performance and computational efficiency.

The significance of this work lies in exploring the capability of deep learning and Transfer Learning techniques in medical image classification tasks. The obtained results demonstrate the effectiveness

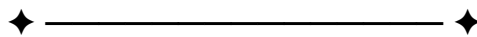
of pretrained CNN architectures for MRI brain tumor classification and highlight the importance of selecting appropriate training parameters to achieve reliable classification performance.

This dissertation is organized into three main chapters :

- The first chapter introduces the problem statement and presents the fundamental concepts related to medical image classification.
- The second chapter provides the theoretical background of Artificial Intelligence, Deep Learning, and Transfer Learning techniques.
- The third chapter presents the proposed methodology, including data preprocessing, data augmentation, model training, hyperparameter optimization, and experimental results analysis.

Chapter 01:

Problem Formulation



1.1 Introduction

Medical image analysis has become one of the most prominent modern research fields due to its pivotal role in improving disease diagnosis and healthcare quality, particularly Magnetic Resonance Imaging (MRI) technology, which is widely used in detecting brain abnormalities thanks to its ability to produce detailed images of brain tissues. However, manual analysis of these images remains time-consuming and heavily dependent on specialized medical expertise. With the rapid development of Artificial Intelligence (AI) and Deep Learning, image classification systems have achieved remarkable performance in automatically analyzing medical images, most notably Convolutional Neural Networks (CNNs), which have demonstrated high efficiency in extracting features and classifying brain MRI images with notable accuracy. Nevertheless, existing challenges such as data scarcity, class imbalance, and limited annotated datasets continue to affect the performance of deep learning models in medical imaging applications. Therefore, this chapter aims to present the fundamental concepts related to medical image classification and brain MRI analysis, highlight the major challenges in this domain, and shed light on the importance of Deep Learning and Transfer Learning techniques in enhancing brain tumor classification performance and improving its accuracy.

1.2 Introduction to image classification:

1.2.1 Definition and importance of image classification :

- **Definition:** The image classification is a classical problem of image processing, computer vision and machine learning fields. Image classification is a complex procedure which relies on different components. [1]
- **Important:** As one of the important research directions in the field of computer vision, image classification has a wide range of applications. Image classification is an important foundation for pattern recognition, machine learning, and artificial intelligence. Image classification generally includes three steps: region of interest selection, feature extraction, and classifier modeling. Among them, feature extraction of images is an important foundation for completing other tasks. [2]

1.3 Architecture of image classification

Figure 1.1 represents the general architecture of image classification, illustrating the sequential stages of the process starting from Datasets and Preprocessing, leading to Feature extraction, and concluding with Classification and Decision.

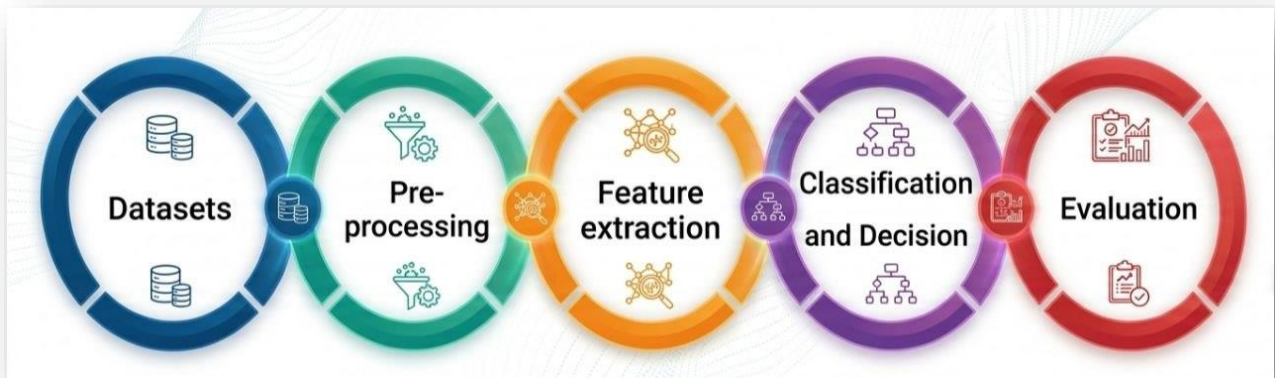


Figure1.1 General architecture of image classification.

1.3.1 Datasets:

A dataset is an organized collection of data samples that can be used for analysis, processing, and different computational or statistical applications. Each sample consists of a set of features and may include a label in supervised learning. In machine learning and deep learning, the performance and generalization capability of models are highly dependent on the quality, quantity, and diversity of this training data. However, acquiring largescale, diverse datasets in realworld applications presents significant challenges, particularly in specialized domains such as medical imaging, autonomous systems, and remote sensing. These challenges include [3]:

➤ Medical imaging :

- **Data Scarcity:** Limited availability of annotated data due to privacy concerns, ethical restrictions, and the need for expert labelling, especially in medical and healthcare domains [3].
- **Class Imbalance :** Unequal representation of classes within datasets leads to biased learning, causing models to underperform on minority or rare classes.
- **High Cost of Data Collection and Annotation :** Acquiring and labelling highquality datasets is often timeconsuming, labourintensive, and expensive.
- **Traditional data augmentation :** methods such as rotation, flipping, and scaling provide only modest variability and often fall short in representing the complex, realworld variations that are essential for effective model training. As a result, these constraints can hinder a model's capacity to generalize to new, unseen data, leading to several issues :
 - Decreased classification accuracy, particularly for underrepresented classes.
 - Increased risk of overfitting due to insufficient or imbalanced training data.
 - Dependence on extensive manual intervention or costly data acquisition pipelines [3].

1.3.2 Preprocessing:

Preprocessing refers to transformations applied to raw data before it is used for training a model, in order to improve learning efficiency and model performance [4].

➤ **Examples include**

- Normalization.
- Standardization.
- Rescaling.
- Noise removal.

1.3.3 Feature extraction :

Feature extraction is a fundamental process in machine learning and deep learning that aims to transform raw data into a reduced set of meaningful and informative features capable of representing the essential characteristics of the data while minimizing redundancy and complexity. This process enables learning models to focus on the most relevant information, thereby improving computational efficiency, model performance, and generalization capability. [4]

Traditionally, feature extraction has relied on handcrafted techniques such as SIFT, HOG, and LBP, where features are manually designed based on prior domain knowledge and extracted before the training stage. Although these conventional methods are computationally efficient and relatively simple to implement, they often struggle to capture complex patterns and highdimensional representations. In contrast, deep learning approaches, particularly Convolutional Neural Networks (CNNs), perform automatic feature extraction directly from raw data through multiple hierarchical layers that progressively learn representations ranging from simple lowlevel patterns to complex highlevel features. This hierarchical learning capability has demonstrated superior performance and enhanced generalization in complex applications such as image and video recognition, despite the higher computational cost and the need for largescale datasets. [5]

1.3.4 Classification and Decision :

Classification in deep learning refers to the supervised learning task where a deep neural network learns a mapping function from input data to categorical labels by extracting hierarchical representations and determining the correct class for each input instance [5].

1.3.5 Evaluation:

Evaluating model performance is essential to ensure their reliability and clinical applicability. This phase involves the use of appropriate metrics and rigorous validation techniques.

➤ Evaluation Metrics:

Several metrics are commonly used to evaluate classification models:

- **Accuracy:** Proportion of correctly classified examples. However, it can be misleading in cases of class imbalance

$$Accuracy = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \dots \dots \dots (1)$$

The accuracy of a classification model is calculated using the following equation :

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \dots \dots \dots (2)$$

Where:

TP = True Positives (correctly predicted positive cases)

TN = True Negatives (correctly predicted negative cases)

FP = False Positives (incorrectly predicted as positive)

FN = False Negatives (incorrectly predicted as negative)

- **Sensitivity:** Model's ability to detect positive cases. Critical in medical diagnosis to minimize false negatives

$$Sensitivity = \frac{TP}{(TP + FN)} \dots \dots \dots (3)$$

TP = True Positives (correctly predicted positives)

FN = False Negatives (incorrectly predicted as negative)

- **Specificity:** Model's ability to identify negative cases. Important to avoid false positives.

$$Recall = \frac{Tp}{Tp+FP} \dots \dots \dots (4)$$

TN = True Negatives (correctly predicted negatives)

FP = False Positives (incorrectly predicted as positive)

F1 Score: Harmonic mean between precision and sensitivity, useful for balancing these two criteria:

$$F1 \text{ Score} = \frac{(2 \cdot Precision \cdot Sensitivity)}{(Precision + Sensitivity)} \dots \dots \dots (5)$$

Where:

$$Precision = \frac{TP}{(TP + FP)} \dots \dots \dots (6)$$

Precision = Out of everything the model said was tumor, how many were correct.

Sensitivity = Out of all real tumors, how many the model found.

1.4 Problem Statement and Research Objectives

- **definition of the data scarcity :**

we define data scarcity in my research as the significant lack of high-quality, annotated medical datasets, which hinders the effective training of deep learning models.

1.5 Conclusion:

In this chapter, the fundamental concepts of image classification and medical image analysis were presented. The main stages of image classification systems, including preprocessing, feature extraction, classification, and evaluation, were discussed. In addition, the chapter highlighted the challenges of medical imaging applications and introduced the importance of Deep Learning and Transfer Learning techniques for brain MRI classification.

Chapter 02 :

Theoretical Background and Fundamental Concepts of AI and Learning Methods



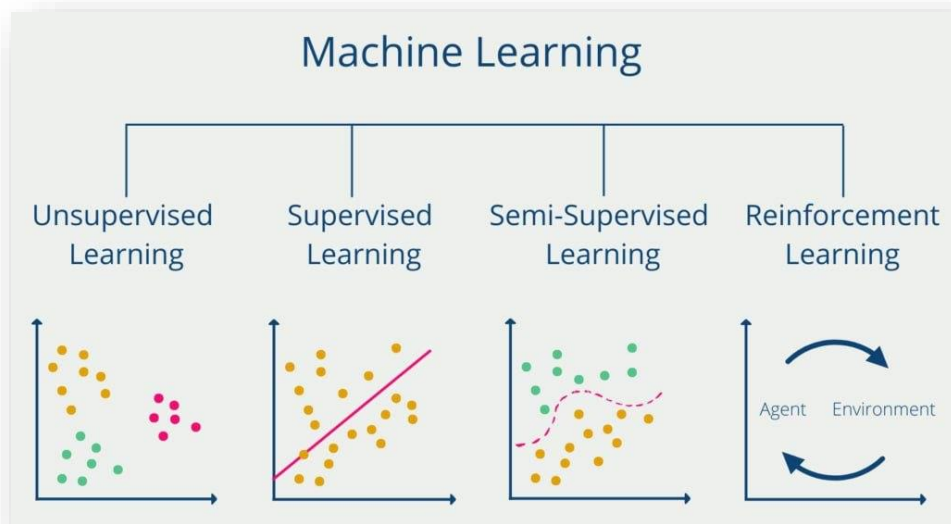
2.1 Introduction

This chapter presents the theoretical background and fundamental concepts related to Artificial Intelligence and learning approaches, aiming to provide a comprehensive overview of the main fields that support the development of intelligent systems, including Machine Learning, Deep Learning, and Transfer Learning, where understanding these concepts is considered essential for building a solid foundation for the proposed work, as they represent the core techniques used in modern image processing and classification systems. Accordingly, the main objective of this chapter is to introduce and explain the key concepts of AI and learning methods relevant to this study, by clarifying the differences and relationships between Machine Learning, Deep Learning, and Transfer Learning, and providing the reader with the necessary theoretical knowledge that will support the methodology and implementation presented in the following chapters.

2.2 Definition of AI :

Artificial Intelligence is defined as the scientific and engineering field concerned with developing computer systems and software capable of simulating human mental processes, such as reasoning, learning, and problemsolving, through data processing and the construction of algorithms that enable the system to respond to inputs in an intelligent manner [6].

2.3 Machine learning



[7]Figure 2.1 An illustrative diagram of Machine Learning

2.3.1definition

Machine learning is considered one of the fundamental branches of artificial intelligence, as it relies on a set of algorithms trained on large and diverse datasets with the aim of building selflearning models capable of extracting patterns, thereby enabling them to predict outcomes and classify information without the need for direct human intervention. Machine learning has demonstrated a broad presence across various contemporary commercial applications, most notably in product recommendation systems based on users' purchase histories, predictive models for stock market fluctuations, and automated text translation systems between different languages [7].

2.3.2Types of machine learning

Machine learning encompasses several distinct types, as shown in the Figure 2.1 each of which plays a fundamental role in powering the wide array of digital products and services that have become integral to modern daily life. Although these various types share a common overarching objective namely, the development of systems and applications capable of operating autonomously without human intervention they differ considerably in terms of the specific methodologies and approaches they employ to achieve this goal.

In order to provide a clearer understanding of how these types are distinguished from one another, the following section presents a comprehensive overview of the four primary categories of machine learning that are predominantly utilized in contemporary applications.

1. Supervised machine learning

Supervised learning is one of the most widely used approaches in machine learning, where algorithms are trained using labeled datasets that contain predefined correct answers. During the training process, the model learns the relationship between input data and their corresponding labels, enabling it to recognize patterns and make accurate predictions on new unseen data. For example, an algorithm can be trained on flower images labeled by type, allowing it to classify new flower images correctly. Due to its accuracy and reliability, supervised learning is extensively used in predictive analysis and data classification applications [7]. The systematic data flow and mapping mechanism from labeled training inputs to accurate predictions are illustrated in Figure 2.2.

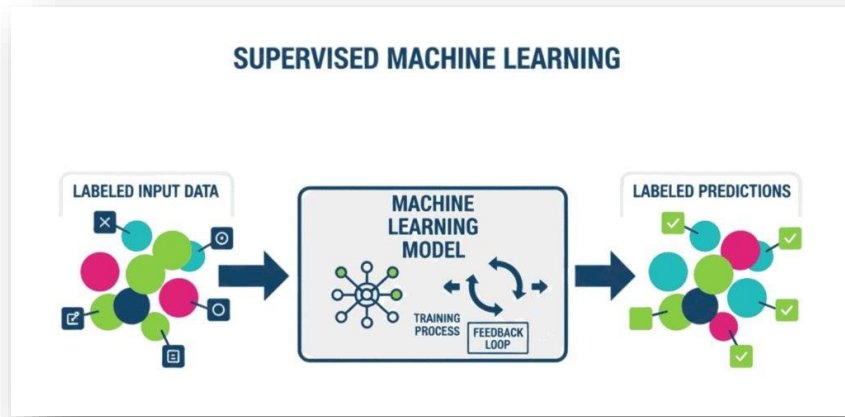


Figure 2.2: Schematic representation of data flow from labeled inputs to predictions

2. Unsupervised machine learning

Unsupervised learning is a machine learning approach that relies on unlabeled data for training. In this method, algorithms are not provided with predefined answers or categories, but instead learn to identify hidden patterns, structures, and relationships within the data independently. For example, an algorithm can analyze large amounts of unlabeled social media user data to discover common behavioral trends and interaction patterns. Due to its ability to process large volumes of unstructured data efficiently, unsupervised learning is widely used for pattern discovery, data analysis, and extracting meaningful insights from complex datasets [7]. The process of feature extraction and dimensionality reduction for complex datasets is shown in Figure 2.3.

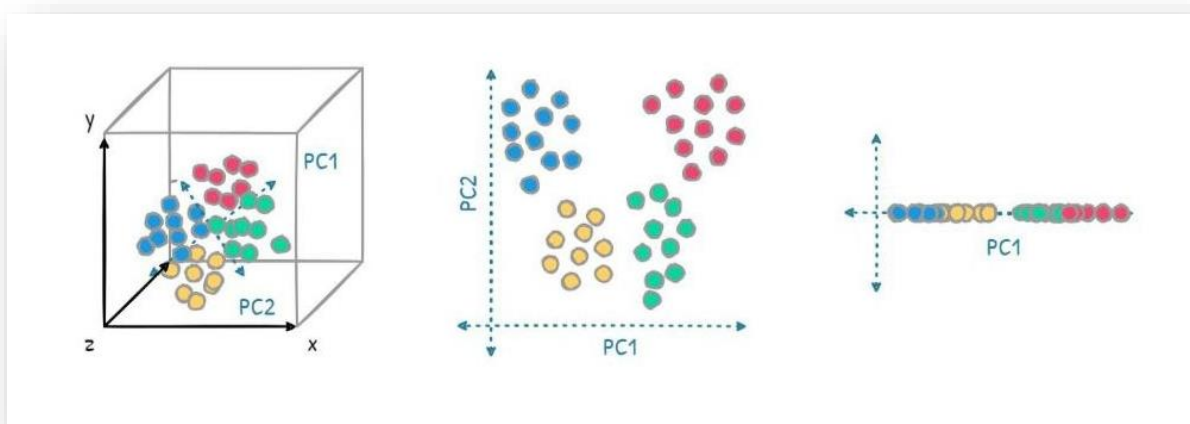


Figure2.3: Dimensionality reduction technique as part of Unsupervised Machine Learning. [7]

3. Semisupervised machine learning

Semisupervised learning is a hybrid machine learning approach that combines both supervised and unsupervised learning techniques by using labeled and unlabeled data during training. Typically, the algorithm is first trained on a small amount of labeled data to establish initial guidance, and then it learns from a much larger set of unlabeled data to improve its performance. For example, a model may begin training with a limited number of labeled speech samples and later use a large collection of unlabeled speech data to develop an accurate speech recognition system. Because it reduces the need for large labeled datasets while still achieving strong performance, semisupervised learning is widely used in classification and prediction tasks where labeled data is scarce or expensive to obtain [7]. The workflow and mechanism of this hybrid learning approach, showcasing how labeled and unlabeled data are integrated, are illustrated in Figure 2.4.

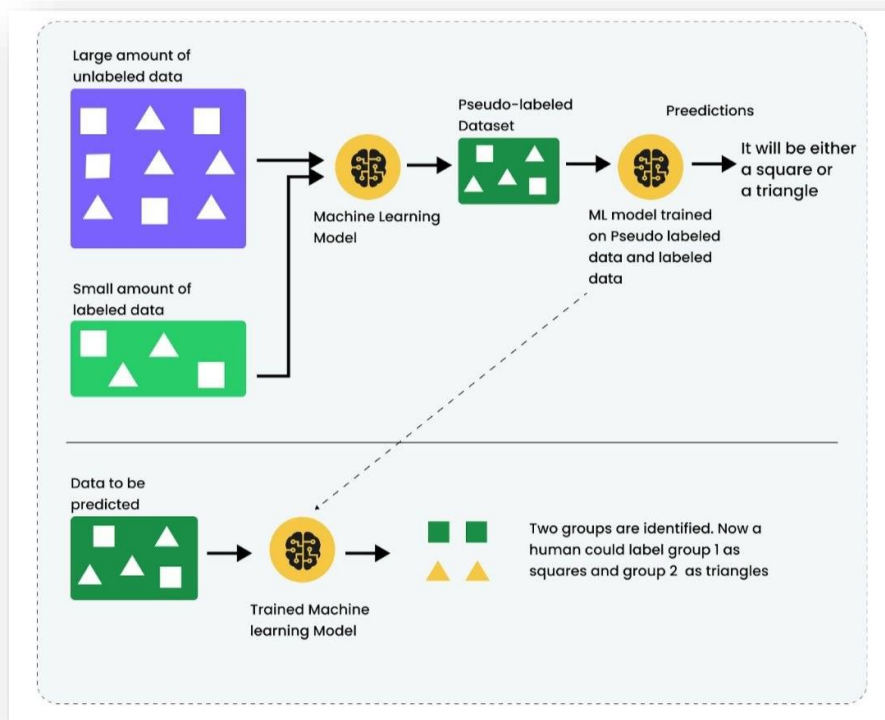


Figure2.4: Schematic diagram of the Semisupervised learning process

4. Reinforcement learning

Reinforcement learning is a type of machine learning where algorithms learn through interaction with an environment by trial and error. During training, the system performs actions and receives feedback based on the results of those actions. Over time, it gradually builds an understanding of the environment and improves its behavior in order to maximize desired outcomes, similar to how a child learns from experience.

For example, an algorithm can be trained by repeatedly playing games of chess, where it learns from both its wins and losses to refine its strategy. Reinforcement learning is commonly applied in situations that require a sequence of decisions, such as game playing or text summarization, where the model must continuously choose the most appropriate actions to achieve a specific goal [7]. The following figure 2.5 serves as a prime example.

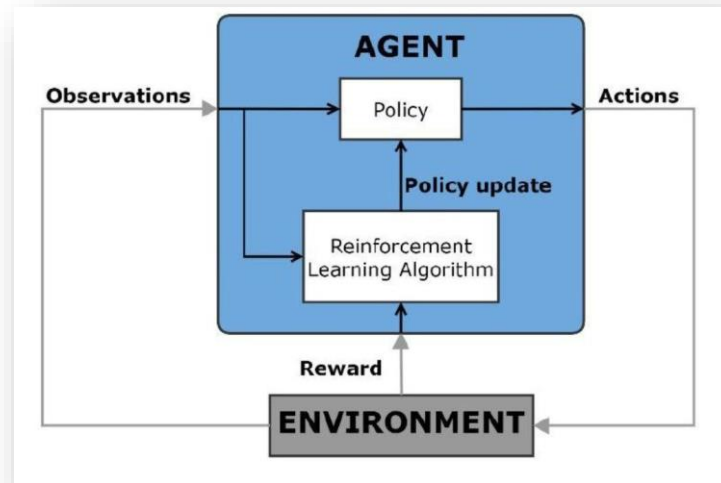


Figure 2.5The architectural diagram of the Reinforcement Learning system.

2.3.3 Examples and use cases

Machine learning is typically the most mainstream type of AI technology in use today. Some of the most common examples of machine learning that you may have interacted with in your daytoday life include :

- Recommendation engines that suggest. products, songs, or television shows to you, such as those found on Amazon, Spotify, or Netflix.
- Speech recognition software that allows you . .to convert voice memos into text.
- A bank's fraud detection services. .automatically flag suspicious transactions.
- Selfdriving cars and driver assistance. features, such as blindspot detection and automatic stopping, to improve overall vehicle safety [7].

2.3.4How does machine learning ?work

Machine Learning works by using algorithms that are trained on large sets of data to recognize patterns and make predictions or classifications on new data. During the training process, the algorithm learns from labeled examples, such as images or numerical data, and adjusts its internal parameters to improve accuracy.

Once training is completed, the algorithm becomes a machine learning model capable of performing specific tasks such as image recognition, prediction, or decisionmaking. In more advanced systems, multiple models can be combined in layered structures to form deep learning networks, which are used for more complex tasks like language processing and chatbot systems .

2.4 Deep Learning

2.4.1 definition

Deep Learning is a modern branch of artificial intelligence that uses multilayered neural networks to automatically learn hierarchical representations from data. Unlike traditional methods, it eliminates the need for manual feature extraction by learning directly from large datasets. This enables it to model complex patterns and achieve high performance in tasks such as image classification, natural language processing, and medical analysis [8].

2.4.2Types of Deep Learning

Deep learning encompasses a variety of architectures, each designed to address specific types of data and learning tasks, making it a versatile and powerful approach in modern artificial intelligence.

1. Convolutional Neural Networks (CNN)

The convolution operation represents the core component of CNN architectures, where learnable filters are applied to the input image to generate feature maps capable of extracting important visual patterns such as edges, textures, and shapes. This operation can be expressed as follows :

$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(i - m, j - n)K(m, n).....(7)$$

Where:

- **I** represents the input image.
- **K** represents the convolution kernel (filter).
- **S** represents the generated feature map.
- **(i,j)** represent spatial positions in the image.

- ❖ After the convolution stage, activation functions such as the Rectified Linear Unit (ReLU) are applied to introduce nonlinearity into the network and improve learning capability :

$$f(x) = \max(0, x).....(8)$$

Where:

- **f(x)** represents the output of the activation function.

- $\mathbf{x}$ represents the input value to the neuron.
- ❖ To reduce the spatial dimensions of feature maps and decrease computational complexity, pooling operations are employed. Max Pooling selects the maximum value within a local region of the feature map, which helps preserve important features while reducing overfitting :

$$p(i, j) = \max(s(m, n)) \dots \dots \dots (9)$$

Where:

- $\mathbf{P(i,j)}$ represents the pooled feature map.
- $\mathbf{S(m,n)}$ represents values inside the pooling region
- ❖ At the final stage of classification, the Softmax function converts network outputs into probability distributions corresponding to the target classes :

$$p(y_i) = e^{(z_i)} / \sum e^{z_j} \dots \dots \dots (10)$$

Where:

- $\mathbf{P(y_i)}$ represents the probability of class i.
- $\mathbf{z_i}$ represents the output score for class i.
- $\mathbf{\sum e^{(z_j)}}$ represents the sum of exponential scores for all classes.

The performance of the CNN model during training is evaluated using the CrossEntropy loss function, which measures the difference between predicted probabilities and true labels :

$$L = - \sum y_i \log(\hat{y}_i) \dots \dots \dots (11)$$

Where:

- $\mathbf{L}$ represents the loss value.
- $\mathbf{y_i}$ represents the true label.
- $\mathbf{\hat{y}_i}$ represents the predicted probability.

- ❖ In Transfer Learning, pretrained CNN models are adapted to a target medical dataset through a finetuning process, where pretrained weights are updated according to the new task requirements :

$$W_{new} = w_{pretrained} + \Delta w \dots \dots \dots (12)$$

Where:

w_{new} represents the updated weights.

$w_{pretrained}$ represents the pretrained model weights.

ΔW represents the weight updates during finetuning.

- ❖ During the optimization process, network weights are iteratively updated using optimization algorithms such as Gradient Descent or Adam Optimizer in order to minimize the loss function and improve classification performance :

$$w(t + 1) = W_t - \eta \nabla L(w_l) \dots \dots \dots (13)$$

Where:

- w_t represents the current weights.
 - $W(t+1)$: represents the updated weights.
 - η : represents the learning rate.
- $\nabla L(w_l)$: represents the gradient of the loss function.

The fundamental operations involved in CNN architectures and the overall training workflow are illustrated in Figure 2.6 .

CNN Operations and Training Workflow

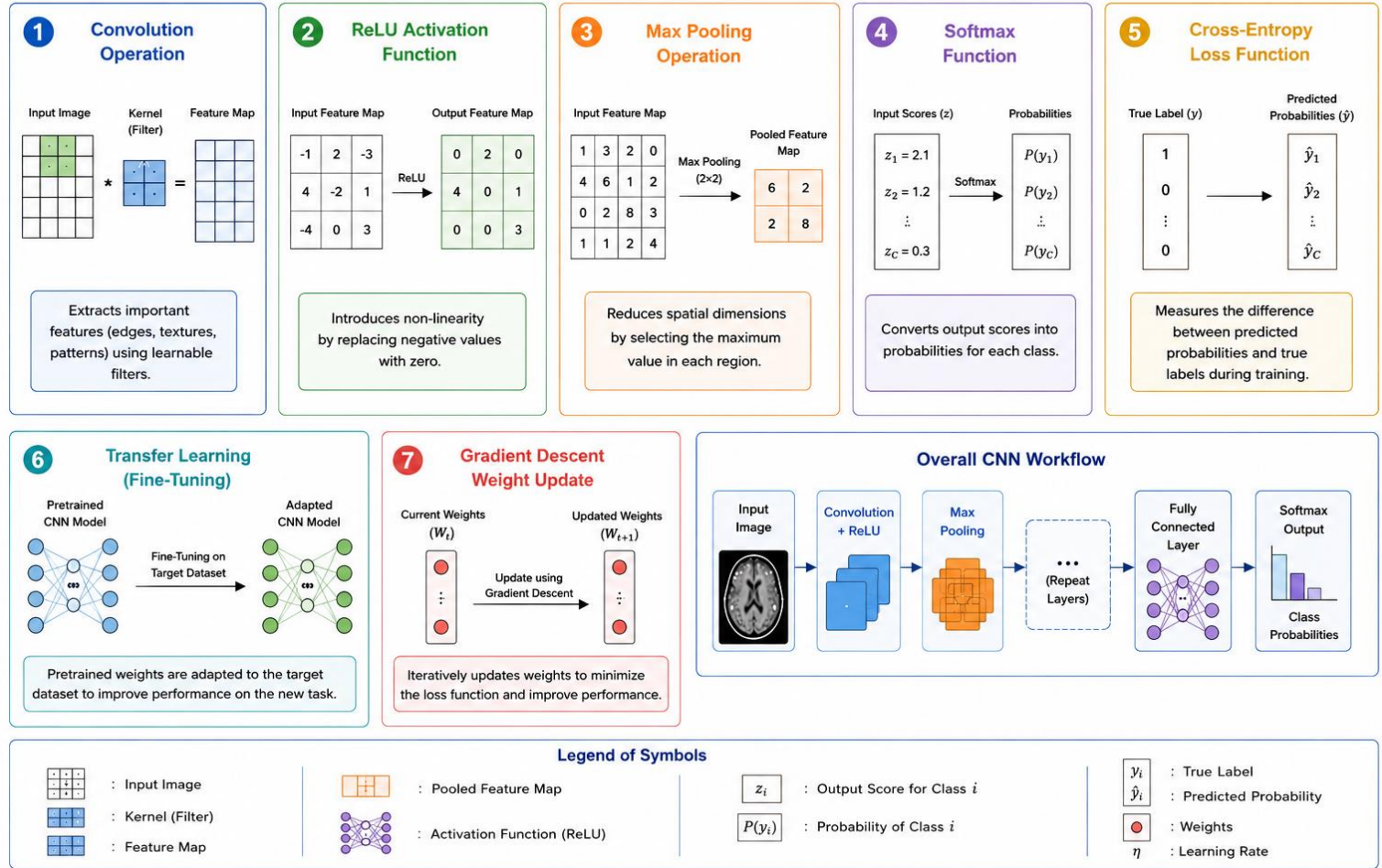


Figure 2.6: Fundamental CNN Operations and Training Workflow for Deep LearningBased Image Classification.

CNN architectures have evolved into various types, each designed with specific characteristics and advantages to address different tasks and improve performance.

a. ResNet18

ResNet18 is a convolutional neural network composed of 18 layers that uses residual learning (skip connections) to solve the vanishing gradient problem. It allows the network to learn deeper representations while maintaining stable training.

b. MobileNetV2

MobileNetV2 is a lightweight CNN designed for mobile and embedded devices. It uses depthwise separable convolutions and inverted residual blocks to reduce computational cost while maintaining good accuracy.

c. DenseNet201

DenseNet201 is a deep CNN architecture with 201 layers where each layer is connected to every other layer in a feedforward manner. This dense connectivity improves feature reuse and reduces the vanishing gradient problem.

d. ResNet50

ResNet50 is a 50layer deep convolutional neural network that uses residual connections to enable efficient training of very deep networks. It is widely used in image classification and transfer learning tasks due to its strong performance.

2. Recurrent Neural Networks (RNN)

Recurrent Neural Networks are designed to handle sequential and temporal data by introducing memory through feedback connections. This allows the network to retain information from previous inputs and use it to influence current predictions. RNNs are widely used in natural language processing, speech recognition, and timeseries forecasting [9]. The structural implementation of skip connections or residual links within the network architecture is illustrated in Figure 2.7.

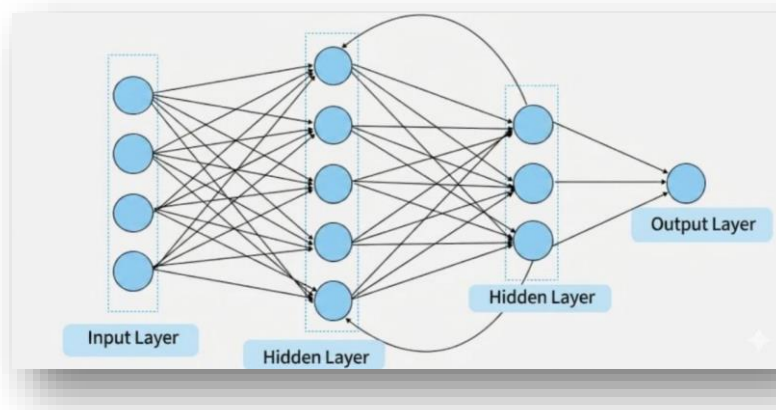


Figure 2.7: Recurrent Neural Network with Skip Connections / Residual Connections [9]

3. Deep Neural Networks (DNN)

Deep Neural Networks are artificial neural networks composed of multiple hidden layers between the input and output layers. These networks are capable of learning complex nonlinear relationships in data through hierarchical feature extraction. DNNs are used in various applications such as classification, prediction, and pattern recognition [10]. The hierarchical architecture of a deep neural network, including the input, output, and multiple hidden layers, is illustrated in Figure 2.8.

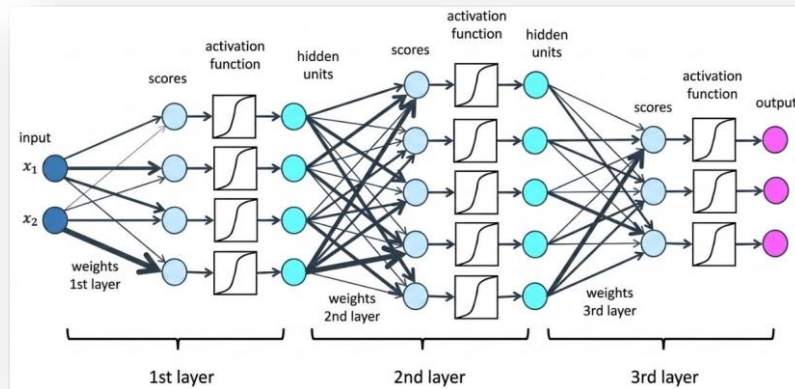


Figure 2.8 Multilayer Feedforward Neural Network Diagram [10]

4. Generative Adversarial Networks (GAN)

Generative Adversarial Networks consist of two competing neural networks : a generator and a discriminator. The generator creates synthetic data, while the discriminator evaluates its authenticity. Through this adversarial process, GANs learn to generate highly realistic data, especially in image synthesis and data augmentation [11]. The detailed architecture and data flow of a Generative Adversarial Network (GAN) are shown in Figure 2.9.

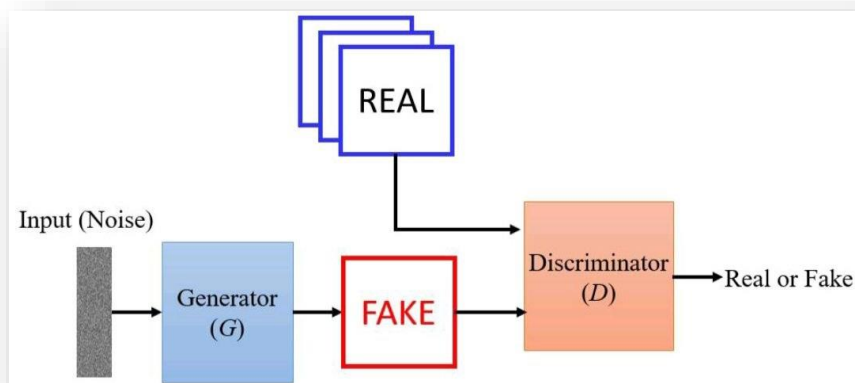


Figure 2.9 : Generative Adversarial Networks GAN Architecture [12]

2.5 Transfer Learning :

2.5.1 definition of transfer learning

Transfer Learning aims to improve the performance of models in target domains by transferring knowledge from related source domains, reducing the need for large amounts of targetdomain data. Due to its wide applications, this field has become popular and promising in machine learning [13].

2.5.2 type of transfer learning

However, previous studies have addressed Transfer Learning methods in a fragmented way and have not fully covered recent advancements:

1. Inductive Transfer Learning

In this type, the source and target tasks are different, but the knowledge learned from the source task helps improve learning in the target task.

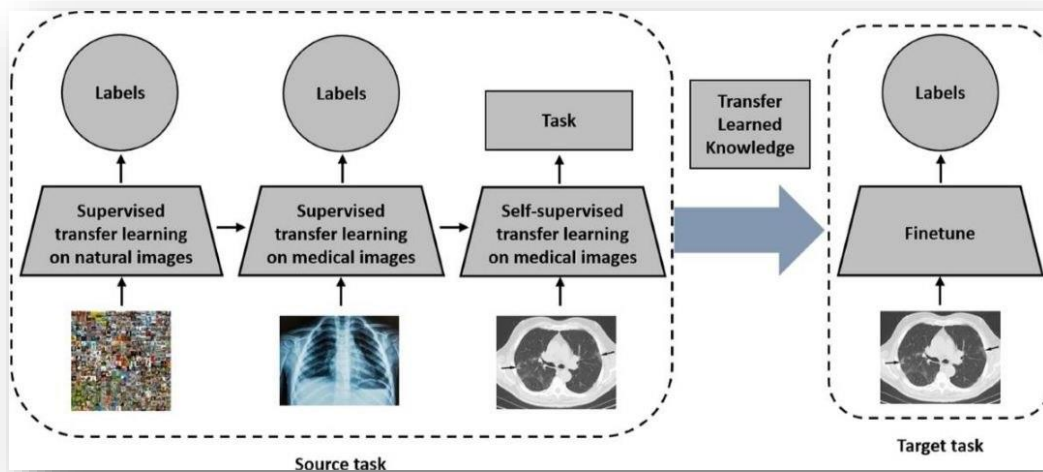


Figure 110: Transfer Learning Process Diagram

The figure2.10 represents the transfer of knowledge from one task to a different task with the presence of supporting data.

2. Transductive Transfer Learning

In this case, the source and target tasks are the same, but the domains are different (different data distributions).

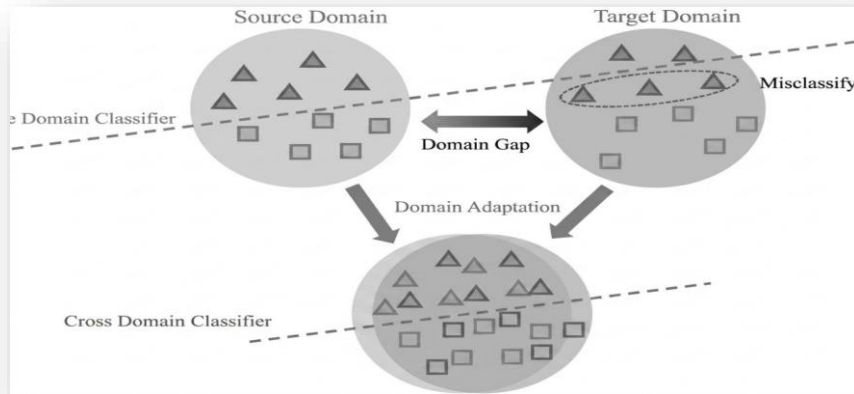


Figure2.11 Domain Adaptation Process

3. Unsupervised Transfer Learning

This type is used when there are no labeled data in either the source or target domain. It focuses on transferring knowledge for feature extraction or clustering tasks, as shown in Figure 2.12.

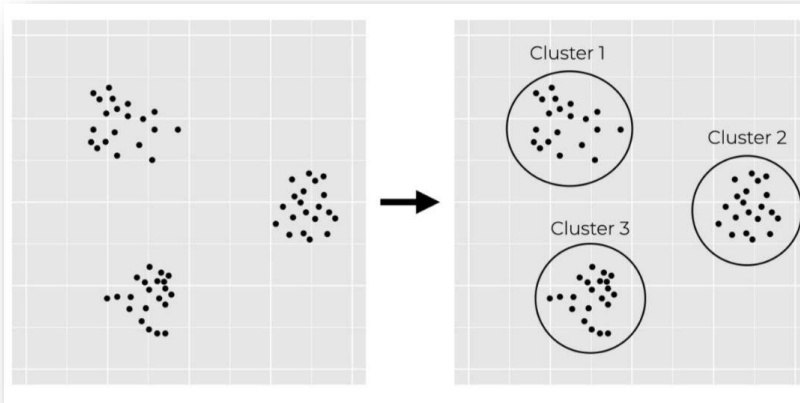


Figure 2.12: Discover patterns without using labels

2.6 Why Transfer Learning in Medical Image Classification:

Transfer Learning is a deep learning technique that allows the reuse of pretrained models for new classification tasks. Instead of training a neural network from scratch, models such as ResNet, VGG16, and InceptionV3 are adapted to medical imaging applications. These models are initially trained on large datasets and then finetuned using medical images. In medical image classification, transfer learning is particularly important because medical datasets are usually limited and expensive to annotate. Pretrained models can effectively extract important visual features from medical images, which improves classification accuracy while reducing training time and computational cost.

This technique is widely used in several medical applications such as brain tumor detection, lung disease diagnosis, breast cancer classification, and diabetic retinopathy analysis. In addition, transfer learning helps reduce overfitting and improves the generalization performance of deep learning models [14].

2.7 State of the art

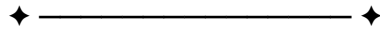
The state of the art in medical image classification has witnessed a significant shift from traditional machine learning methods to deep learning architectures, particularly Convolutional Neural Networks (CNNs). Recent advancements emphasize the efficiency of Transfer Learning (TL) in addressing the scarcity of labeled medical datasets. By leveraging pretrained models such as ResNet, VGG16, and DenseNet, researchers have achieved superior accuracy levels in detecting brain tumors compared to models trained from scratch. These architectures excel at extracting complex hierarchical features from MRI scans, providing a robust foundation for automated diagnostic systems [15].

2.8 Conclusion

This chapter presented the theoretical foundations of Artificial Intelligence, Machine Learning, and Deep Learning, with a particular focus on Convolutional Neural Networks (CNNs) and Transfer Learning techniques. It also discussed the main CNN architectures used in this work, including ResNet18, ResNet50, MobileNetV2, and DenseNet201, highlighting their characteristics and importance in medical image classification tasks.

Chapter 03.

Training protocol and hyperparametre optimization



3.1 Introduction

This chapter presents the implementation and experimental investigation of a Deep Learningbased framework for brain MRI image classification using Transfer Learning techniques. The proposed system was developed and implemented in the MATLAB environment using several pretrained Convolutional Neural Network (CNN) architectures, including ResNet18, ResNet50, MobileNetV2, and DenseNet201. The objective of this chapter is to analyze and compare the performance of these pretrained models for brain tumor classification using MRI images.

The chapter describes the complete workflow of the proposed methodology, starting from dataset preparation and preprocessing to data augmentation, Transfer Learning adaptation, model training, and performance evaluation. In addition, different training hyperparameters, including learning rate, minibatch size, and number of epochs, were investigated in order to study their impact on model convergence, classification accuracy, and generalization capability. Finally, the experimental results obtained from the different CNN architectures are presented and discussed using several evaluation metrics, including accuracy, precision, recall, F1score, and confusion matrix analysis, in order to identify the most effective model for brain MRI classification tasks.

3.2 Proposed Methodology

3.2.1 System Architecture

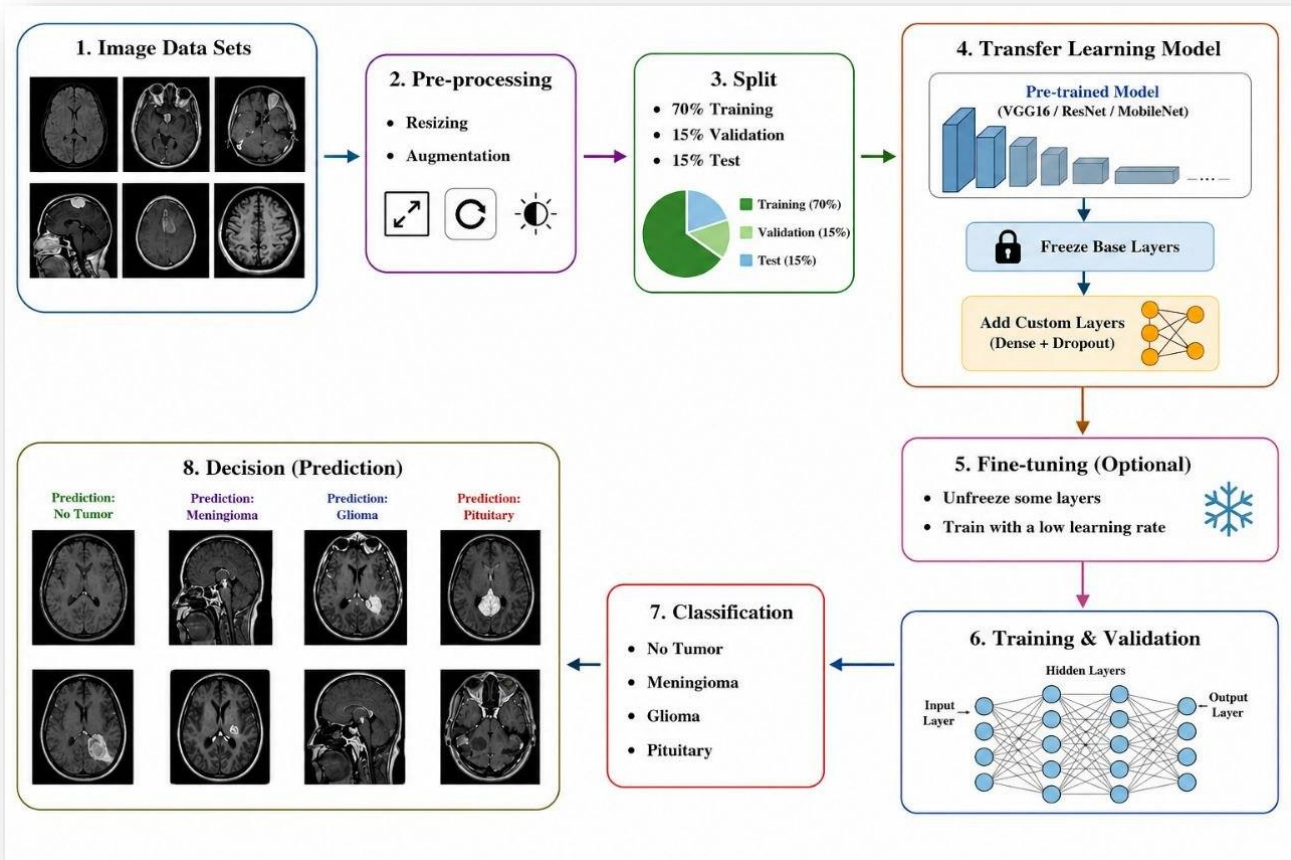


Figure 3.1 Workflow of the proposed brain tumor classification system

Figure 3.1 illustrates a comprehensive and structured pipeline for brain MRI image classification using deep learning techniques, specifically transfer learning. It presents the sequential stages of the process, starting from data preparation and ending with the final prediction.

- The first stage involves the Image Dataset, which consists of brain MRI images categorized into multiple classes such as Glioma, Meningioma, Pituitary tumor, and No Tumor. This dataset forms the foundation for training and evaluating the model.

- Next is the Preprocessing stage, where the data is prepared for model input. This includes resizing images to a uniform dimension and applying data augmentation techniques such as rotation, flipping, and brightness adjustment. These operations enhance data diversity and help reduce overfitting, thereby improving the model's generalization capability.

- In the next stage, we divided the dataset into three parts: 70% for training, 15% for validation, and 15% for testing in order to properly train and evaluate the model. We also used the `splitEachLabel` function with randomization to ensure a balanced distribution of data across all classes

- The fourth stage introduces the Transfer Learning Model, where a pretrained convolutional neural network (e.g., VGG16, ResNet, or MobileNet) is employed. The base layers of the model are frozen to retain previously learned features, and custom layers (such as fully connected Dense layers and Dropout) are added to adapt the model to the specific classification task.

- The fifth stage is finetuning, where some of the previously frozen layers are unfrozen and retrained using a low learning rate. This step allows the model to better adjust to the specific characteristics of the dataset, leading to improved performance.

- Following this, the training and validation phase is conducted. During this stage, the model learns by updating its weights through optimization algorithms while continuously being evaluated on the validation set to prevent overfitting.

In the seventh stage, the model performs Classification, assigning each input image to one of the predefined categories : No Tumor, Meningioma, Glioma, or Pituitary.

Finally, the decision (prediction) stage presents the model's output, showing the predicted class for each MRI image. This step reflects the practical application of the system in assisting medical diagnosis.

- Overall, this pipeline demonstrates a robust and efficient approach to medical image classification, leveraging the power of transfer learning to achieve high performance even with limited data.

3.2.2 Dataset Description

The dataset consists of 7,023 human brain MRI images, categorized into four classes : glioma tumor (1,621 images), meningioma tumor (1,645 images), no tumor (2,000 images), and pituitary tumor (1,757 images). The dataset is divided into training and testing sets. The training data is further split into: 70% for training, where the images are used to teach the system how to recognize different patterns, and 15% for validation, which is used to assess how well the system is learning during the training process, and 15% for test.

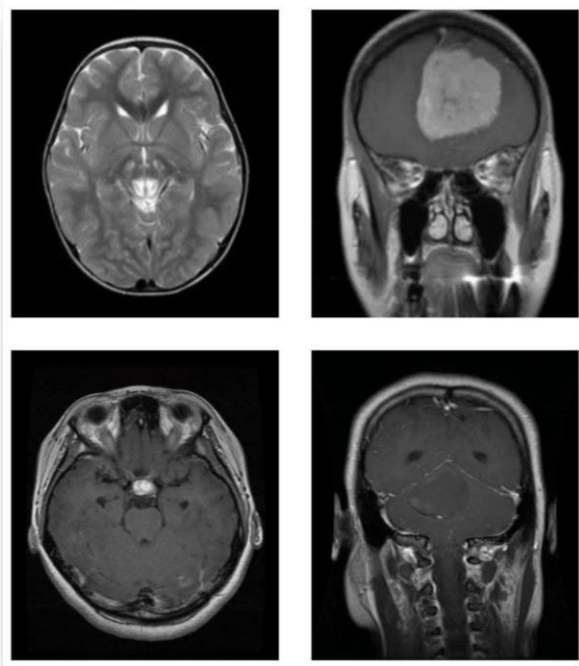


Figure 3.2 :Sample MRI Images from Brain Tumor Dataset

3.2.3 Image Preprocessing and Data Augmentation

➤ Image Resizing:

we programmed the code to automatically extract the input dimensions directly from the input layer of the network. Based on this, all images in the dataset are automatically resized to 224×224 pixels. This procedure ensures that the data fully matches the architectural requirements of the model, which helps maintain processing accuracy and stabilize the training process.

➤ Data Augmentation :

Data Augmentation is a strategy used to increase the amount of data by adding slightly modified copies of already existing data or newly created synthetic data from existing data. It acts as a regularizer and helps reduce overfitting when training a machine learning model. **Figure 3.3** illustrates some common data augmentation techniques applied to MRI images, such as rotation, flipping, zooming, and brightness adjustment, which help improve the diversity and robustness of the dataset.

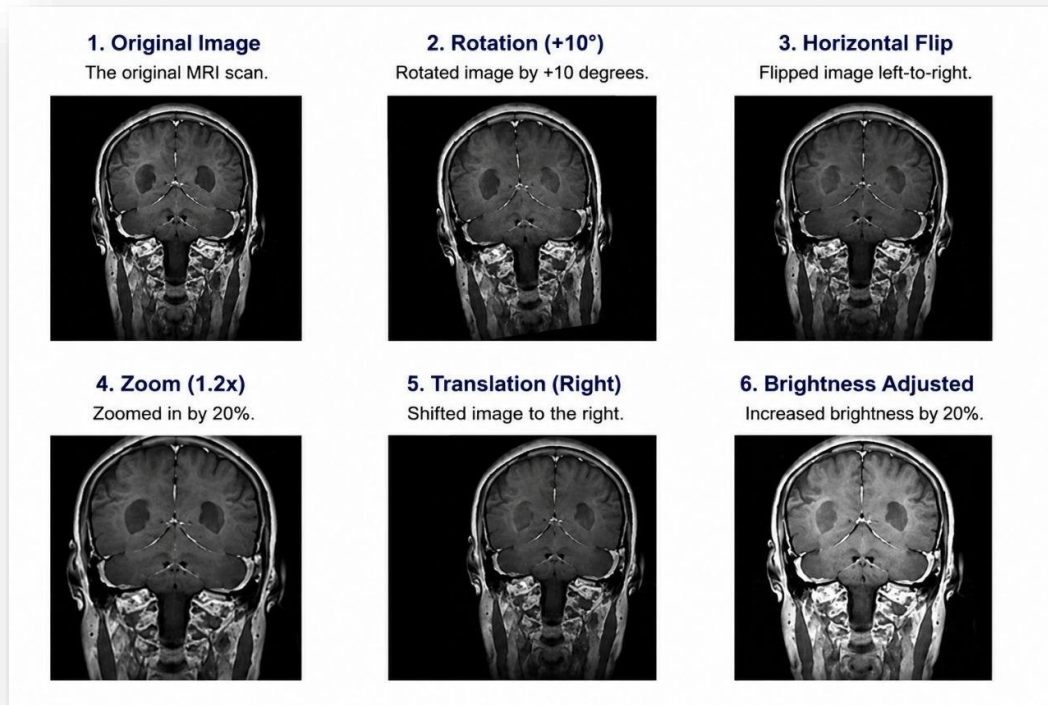


Figure3.3: Examples of RI Data Augmentation

3.2.4 Transfer Learning Models

In this work, we experimented with several Convolutional Neural Network (CNN) architectures instead of relying on a single model. This approach allows us to compare their performance and select the most suitable model for our classification task.

We started with different pretrained CNN models and modified their final layers to adapt them to our dataset. The main architectures used in this study include :

- **ResNet** (Residual Network):

Uses residual connections to solve the vanishing gradient problem, allowing the training of very deep networks.

- **DenseNet** (Dense Convolutional Network):

Connects each layer to every other layer, improving feature reuse and reducing the number of parameters.

- **MobileNet :**

A lightweight architecture designed for efficiency, using depthwise separable convolutions to reduce computation cost.

For each architecture, we applied the following steps. The systematic steps of the transfer learning methodology applied to each CNN architecture are illustrated in Figure 3.4

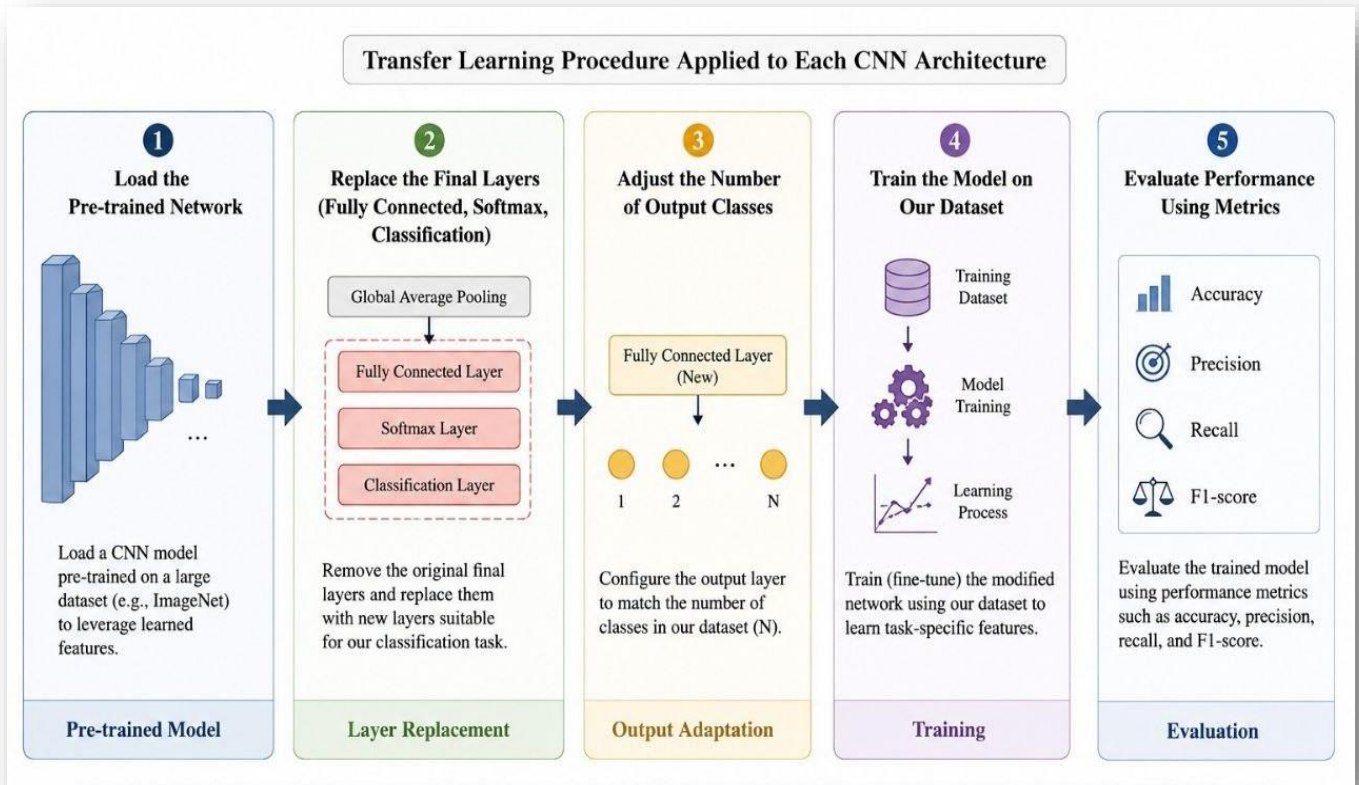


Figure 3.4: Transfer Learning Procedure Applied to Each CNN Architecture

And these are the steps followed:

- Load the pretrained network.
- Replace the final layers (Fully Connected, Softmax, Classification) .
- Adjust the number of output classes.
- Train the model on our dataset.
- Evaluate performance using metrics such as accuracy, precision, recall, and F1score.

3.3 Training and validation

3.3.1 Training option

- In the practical part of this work, we applied the Transfer Learning technique using several Convolutional Neural Network (CNN) models in order to classify brain MRI images with high accuracy. After loading the dataset, we used four pretrained models, namely : DenseNet201, MobileNetV2, ResNet18, and ResNet50.

Next, we modified the final layers of each model so that they match the number of classes in our dataset. In addition, we applied Data Augmentation techniques to increase data diversity, reduce the overfitting problem, and improve the model's generalization ability.

3.3.2 Hyperparameter Optimization

After evaluating the pretrained CNN architectures, the best-performing model was selected for manual hyperparameter tuning. Different training hyperparameters, including the learning rate, mini-batch size, and number of training epochs, were manually adjusted through experimental trials to improve the classification accuracy and the generalization capability of the selected model.

❖ Hyperparameter Optimization Techniques

Hyperparameter optimization is an important step in developing deep learning models, as it aims to determine the set of hyperparameters that provides the best model performance. Several optimization techniques have been proposed in the literature, among which the most widely used are Grid Search, Random Search, and Bayesian Optimization.

➤ **Grid Search:** systematically evaluates all possible combinations of predefined hyperparameter values. Although this method is straightforward and guarantees that every candidate configuration is examined, it becomes computationally expensive as the number of hyperparameters and candidate values increases.

➤ **Random Search:** overcomes this limitation by randomly sampling hyperparameter combinations from a predefined search space. Compared with Grid Search, it generally requires fewer experiments while often achieving similar or even better performance, particularly when only a subset of hyperparameters has a significant impact on the model.

➤ **Bayesian Optimization:** is a more advanced optimization approach that builds a probabilistic model of the objective function and uses information from previous evaluations to intelligently select the next hyperparameter configuration. This strategy usually requires fewer training runs than Grid Search or Random Search while maintaining high optimization efficiency.

In this work, the hyperparameter values were selected through manual hyperparameter tuning based on a series of experimental trials to identify a suitable training configuration for the DenseNet201 model. Therefore, the reported results correspond to manually evaluated hyperparameter configurations rather than an automated hyperparameter optimization process.

Although several automated hyperparameter optimization techniques are available, manual hyperparameter tuning was adopted by evaluating different training hyperparameter configurations through experimental trials to determine the most suitable settings for the DenseNet201 model.

- **Optimizer** : Adam optimizer was used to update the network weights during the training process due to its fast convergence and effective performance in deep learning applications.
- **MiniBatch Size** : The data will be processed in small groups of **16** images per training iteration to maintain stability in memory consumption.
- **Number of Training Epochs** : We have set the model to pass through the entire dataset **10** consecutive times to ensure sufficient learning without reaching the stage of overfitting.
- **Initial Learning Rate** : A low value (10^4) was selected to preserve the weights learned from the pretrained model and to avoid abrupt changes.
- **Data Shuffling** : Images will be reordered randomly at the beginning of each epoch to prevent the network from memorizing the sequence of the data.
- **Validation Data** : An independent dataset will be used to evaluate the model's performance periodically during the training process.
- **Validation Frequency** : The system will perform an accuracy test on the validation data once every **20** training iterations.

Progress Monitoring : We will activate the graphical user interface to monitor the A The specific configurations and training options initialized for the model are detailed in the MATLAB code snippet shown in Figure 3.5. Accuracy and Loss curves in realtime during the learning process.

```
%% 6. Training Options
options = trainingOptions('adam', ...
    'MiniBatchSize', 16, ...
    'MaxEpochs', 10, ...
    'InitialLearnRate', 1e-4, ...
    'Shuffle', 'every-epoch', ...
    'ValidationData', augValidation, ...
    'ValidationFrequency', 20, ...
    'Verbose', false, ...
    'Plots', 'training-progress');
```

Figure 3.5: Training options

3.3.3 Experimental results and discussion :

After completing the training process, each model was tested and evaluated using several performance metrics, including Accuracy, Precision, Recall, and F1score.

Thus, we obtained the results presented in next table :

Table 3.1 Performance comparison of different deep learning models.

model	Test Accuracy	Precision	Recall	F1score
Mobilenet v2	85.46	0.85	0.86	0.85
Resnet 18	95.34	0.95	0.95	0.95
Densnet 201	97.15	0.97	0.97	0.97
Resnet 50	94.39	0.94	0.94	0.94

- The obtained results showed (table 3.1) that the DenseNet201 model achieved the best performance compared to the other models, reaching a test accuracy of 97.15%, with high values of Precision, Recall, and F1score of approximately 0.97. This indicates that the model was able to efficiently extract the important features from brain MRI images.
- The superiority of DenseNet201 can be explained by its architecture, which relies on direct connections between layers. This allows the reuse of extracted features and improves information flow throughout the network, thereby reducing information loss during training. Moreover, this model has a strong ability to learn fine details in medical images, which helped it outperform MobileNetV2, ResNet18, and ResNet50.
- On the other hand, MobileNetV2 achieved the lowest accuracy because it is a lightweight model designed for devices with limited computational resources, making its ability to extract complex features lower than that of deeper models. Meanwhile, ResNet18 and ResNet50 produced good results, but they remained less efficient than DenseNet201 for this classification task.

Based on these results, DenseNet201 can be considered the most suitable and effective model for brain MRI image classification .

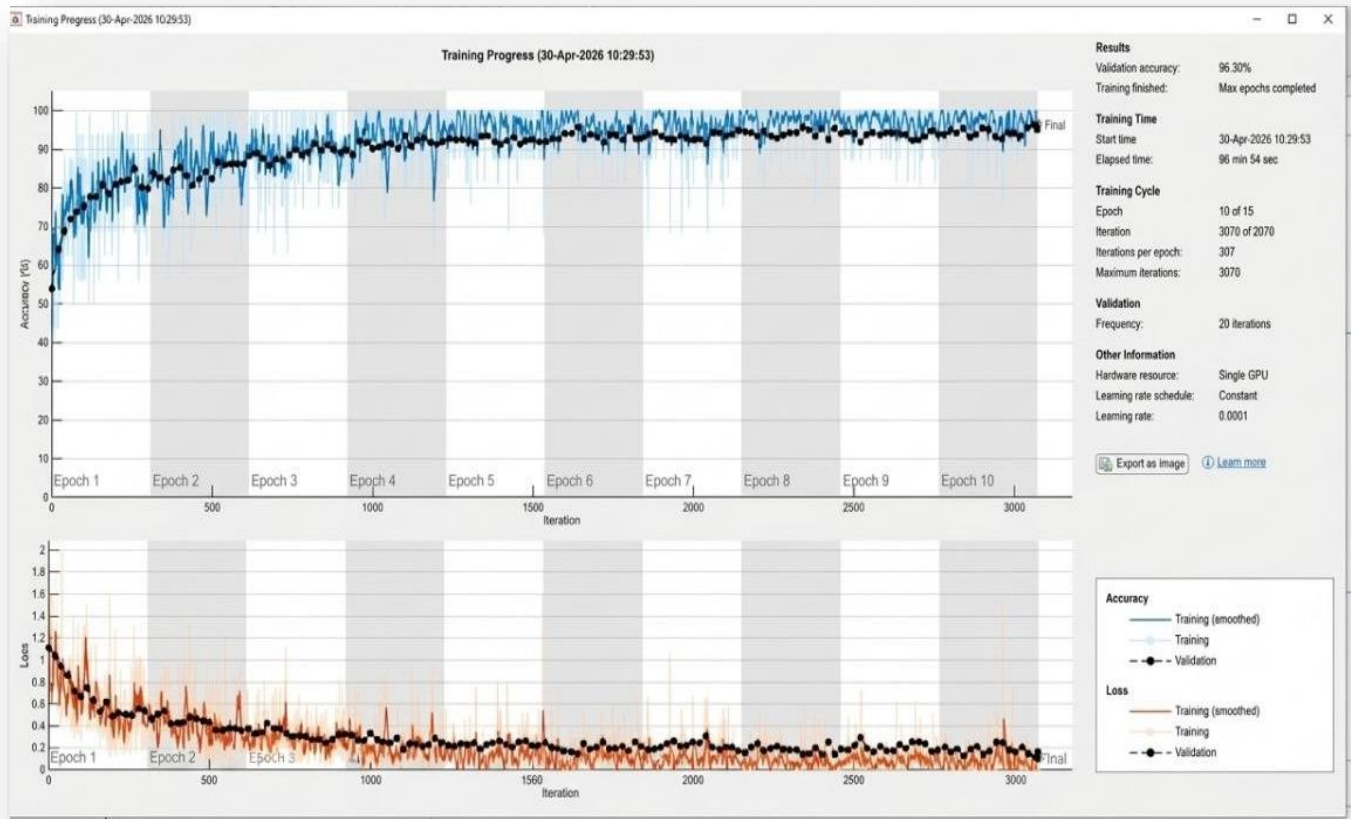


Figure 3.6 Training and validation accuracy and loss curves.

➤ Explanation of the figure

Figure 3.6 illustrates the training and validation curves of the deep learning model over 10 training epochs. The results show that the training accuracy increased steadily and consistently from the early stages, achieving a remarkable validation accuracy of **96.30%**, which indicates the model's strong ability to generalize and effectively learn the underlying patterns in the data. In addition, the loss curves demonstrate a gradual and parallel decrease for both training and validation, with a close convergence between them, confirming the stability of the training process and the absence of overfitting. The training was performed using a Single GPU with a total of 3070 iterations, and the process took approximately 96 minutes and 54 seconds. These results reflect the strong performance of the proposed model and its high efficiency in classification tasks.

- After explaining the training and validation curves, the confusion matrix was used as an additional tool to evaluate the model's performance in more detail. Figure (3.5) presents the confusion matrix for brain tumor classification, showing the number of correct and incorrect predictions for

each class, which helps analyze the model’s accuracy and its ability to distinguish between different tumor types.

Confusion Matrix - Test Data					
True Class	glioma	meningioma	notumor	pituitary	
	glioma	238	5		
	meningioma	6	225	9	6
	notumor		1	299	
	pituitary		3		260
		glioma	meningioma	notumor	pituitary
		Predicted Class			

Figure 3.7 Confusion Matrix for Brain Tumor Classification

The confusion matrix (Figure3.7) illustrates the performance of the model in classifying four types of brain tumor cases. The high values concentrated along the main diagonal, highlighted in blue, indicate very high classification accuracy and a strong agreement between the true and predicted classes. The model demonstrates exceptional performance in identifying the no tumor (notumor) class with 299 correct cases, while the offdiagonal elements show only minimal and limited misclassification, particularly between the meningioma and notumor classes. This further reinforces the statistical reliability of the model in the automated diagnosis of these medical categories.

3.4 Impact of Hyperparameter Optimization

After analyzing the results obtained from the experiment, we found that the DenseNet201 model achieved the best performance compared to the other models, reaching a test accuracy of 97.15%. Therefore, in the second part, we modified only the training settings of this model by changing some parameters such as the number of epochs, MiniBatchSize, and learning rate, with the aim of improving the model’s performance and obtaining better results.

Table 3.2: Comparison of different hyperparameter configurations for DenseNet201

MiniBatchSize	16	10	8
MaxEpochs	10	20	30
frequency	20	20	20
Test Accuracy	98.48	98.86	98.92

➤ **Comment on the table**

- At the beginning of the experiments, we trained the DenseNet201 model using training settings defined by a MiniBatchSize of 16 and MaxEpochs of 10. These settings allowed for a relatively fast training process and reduced training time. However, the limited number of epochs prevented the model from fully exploiting its capacity to learn the finegrained features present in MRI images. In addition, the relatively large MiniBatchSize made the update process more stable, but it may reduce learning precision compared to smaller batch sizes.

- Then, we reduced the MiniBatchSize value from 16 to 8 in order to improve the learning process of the model and enhance its ability to extract fine features from MRI images. This change allowed the model to update its weights more frequently after fewer images, which helped improve classification accuracy and reduce the error rate, although it slightly increased the training time.

- Next, we adjusted the training settings again by setting the MiniBatchSize to 10 and increasing the number of training epochs (MaxEpochs) to 20. This modification allowed the model to train for a longer period and learn more detailed features from the data, which contributed to improving the performance of the DenseNet201 model and achieving better classification results, while maintaining a good balance between training speed and model accuracy.

- After that, we increased the number of training epochs (MaxEpochs) to 30 while keeping the minibatch size (MiniBatchSize) fixed at 10. This adjustment aimed to give the DenseNet201 model a greater opportunity to converge toward the optimal solution by increasing the number of weight updates and allowing multiple passes over the dataset through the network. These more intensive settings helped refine the model's ability to distinguish very subtle differences between classes, which positively reflected on the stability of the accuracy and validation curves. As a result, the model achieved better utilization of the deep architecture, despite the noticeable increase in the training time.

- From these results, we can conclude that achieving the highest levels of accuracy in deep learning models such as DenseNet201 requires a careful balance between learning quality and

computational cost. Although increasing the number of epochs (MaxEpochs) to 30 helped the model reach a state of full convergence and fine refinement of the extracted features, this came at the expense of a noticeable increase in execution time and computational effort. This tradeoff remains acceptable in the context of medical diagnosis, where the highest priority is given to improving accuracy, even by a small margin, as it is a critical factor in the reliability of the system, despite the need to sacrifice training time efficiency.

❖ It has been observed that using a MiniBatchSize of 8 and MaxEpochs of 30 represents the optimal balance point based on the results achieved by the model. This setting helps prevent model overfitting and reduces computational strain. Moreover, using a batch size of 8 instead of 16 or 32 makes the weight updates more frequent, allowing the model to correct its errors more effectively during training.

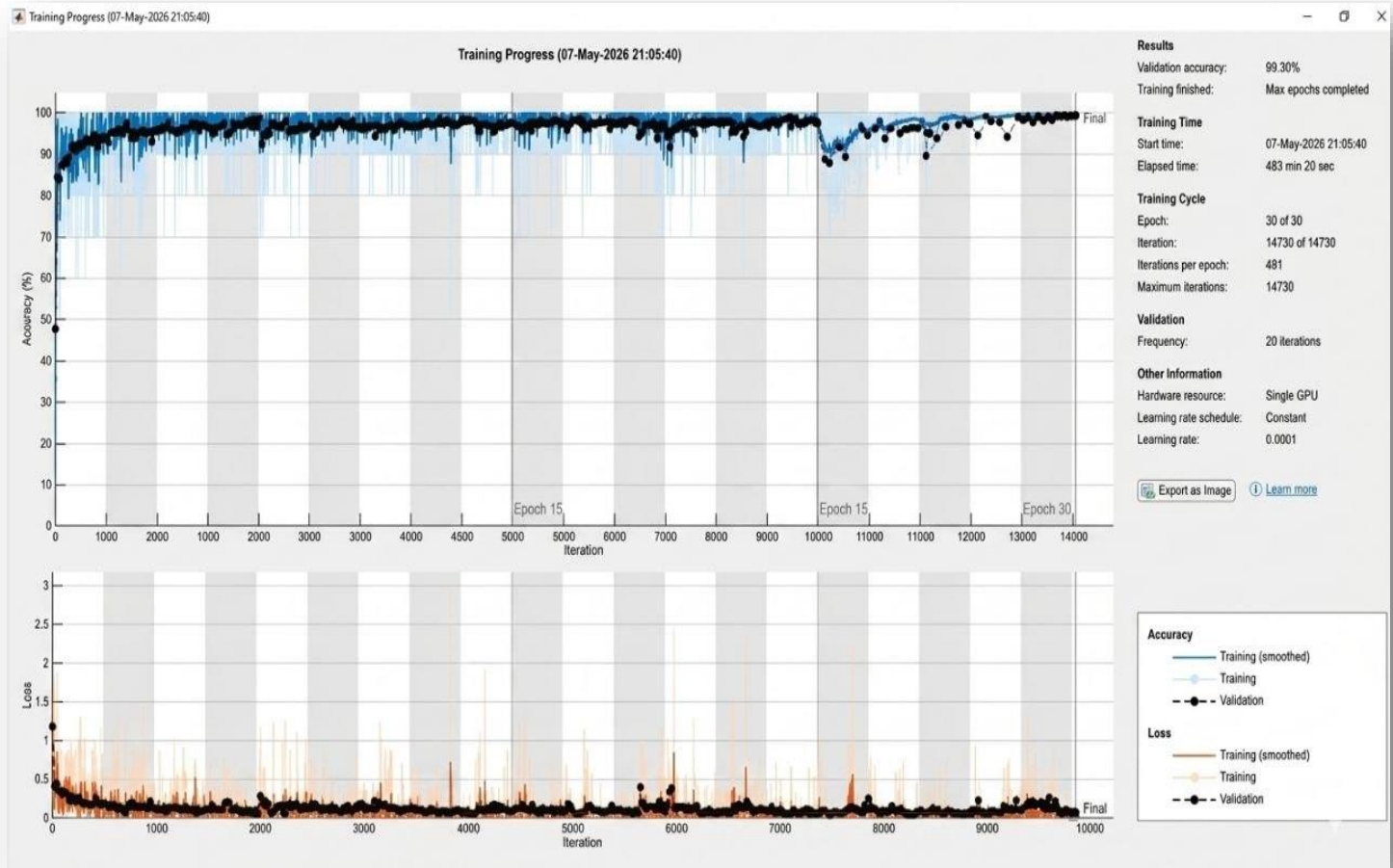


Figure 3.8 Training and validation progress curves for the DenseNet201 model

➤ Explanation of the figure

- The figure 3.8 illustrates the training and validation curves of the deep learning model over 30 training epochs, where the model demonstrated exceptional performance by achieving a validation

accuracy of 99.30%. The graph shows that the training accuracy increased rapidly from the early stages and stabilized at high levels, while the loss curve decreased gradually and consistently until reaching values close to zero. This reflects the model's high learning efficiency and strong generalization capability without suffering from overfitting.

- The training process, which utilized the Adam optimizer with a learning rate of 0.0001 and a total of 14,730 iterations, took approximately 8 hours and 3 minutes. These results confirm that the chosen settings and data preprocessing were optimal for achieving this high level of accuracy and performance stability.

Confusion Matrix - Test Data				
True Class	glioma	meningioma	notumor	pituitary
	glioma	250	1	0
	meningioma		258	0
	notumor		0	305
	pituitary		2	0
Predicted Class				
	glioma	meningioma	notumor	pituitary

Figure 3.9: Confusion matrix of DenseNet201 for brain tumor classification.

➤ Explanation of the figure

The confusion matrix (Figure 3.9) for the test data demonstrates the model's exceptional performance in classifying the four categories, as it successfully identified the vast majority of samples correctly, as shown along the main diagonal of the matrix. The model accurately classified 305 healthy cases (no tumor) and 265 pituitary tumor cases with nearperfect precision, making only one error in each category. It also achieved very strong results in classifying meningioma tumors with 258 correct predictions and glioma tumors with 250 correct predictions, with only a few minor misclassifications that did not exceed 4 cases across the entire test set. These results confirm the

model's high capability to distinguish between closely related classes and minimize confusion among different tumor types to the lowest possible level ..

3.5 Conclusion

This chapter presented the practical implementation and experimental evaluation of the proposed brain MRI classification system using Deep Learning and Transfer Learning techniques. Several pretrained CNN models, including ResNet18, ResNet50, MobileNetV2, and DenseNet201, were trained and evaluated using different

performance metrics such as Accuracy, Precision, Recall, F1score, and Confusion Matrix. The obtained results demonstrated that the DenseNet201 model achieved the best classification performance compared to the other models, especially after optimizing the training parameters and applying preprocessing and data augmentation techniques. The experimental analysis also showed the significant impact of training settings, such as batch size and number of epochs, on improving the model's performance and generalization capability.

Overall, the proposed system achieved highly accurate and reliable results in brain tumor classification, confirming the effectiveness of Deep Learning and Transfer Learning approaches in medical image analysis and their potential to support automated medical diagnosis systems.

Conclusion general

In conclusion, this dissertation presented an applied study that leverages Deep Learning and Transfer Learning techniques for medical image classification, specifically the classification of brain Magnetic Resonance Imaging (MRI) images to detect and differentiate between various types of brain tumors with high accuracy. The obtained results demonstrated the strong potential of these techniques to enhance automated diagnostic systems and reduce full reliance on manual analysis of medical images.

This study relied on several pretrained deep models, namely ResNet18, ResNet50, MobileNetV2, and DenseNet201. A comparative evaluation was conducted using precise performance metrics. The results showed the superiority of the DenseNet201 model in achieving the best accuracy and generalization capability, especially after applying image preprocessing and data augmentation techniques.

The study also highlighted the significant impact of Training Hyperparameter Optimization, where tuning parameters such as minibatch size, number of epochs, and learning rate substantially improved the model's performance and stability. This confirms that the success of deep learning models depends not only on selecting an appropriate architecture but also on carefully configuring the training settings.

In light of the results achieved, several future perspectives can be proposed, including expanding the dataset, experimenting with more recent deep architectures, applying more advanced optimization techniques, and integrating the proposed system into real medical applications used in hospitals and diagnostic centers.

Overall, this dissertation has highlighted the vital role of Artificial Intelligence techniques in the medical field and demonstrated their ability to provide practical and effective solutions to complex medical image analysis problems.

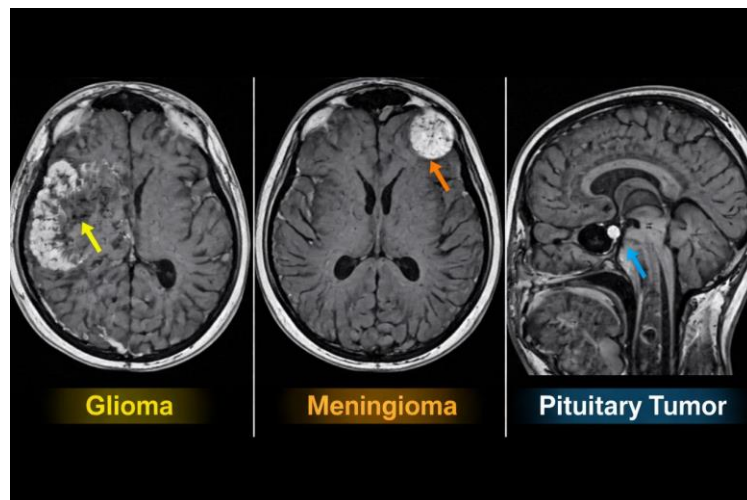
Annex

MRI (Magnetic Resonance Imaging):

Magnetic Resonance Imaging (MRI) is a noninvasive medical imaging technique that uses strong magnetic fields and radio waves to generate detailed images of the internal structures of the body, particularly soft tissues. MRI is widely used in brain analysis because it provides high-resolution images without ionizing radiation, making it essential for detecting abnormalities such as tumors.

Brain Tumor:

A brain tumor is an abnormal growth of cells within the brain or surrounding tissues. Brain tumors can be benign (noncancerous) or malignant (cancerous) and may affect brain function depending on their size, type, and location. Accurate detection and classification of brain tumors using MRI images are crucial for diagnosis, treatment planning, and patient prognosis.



- This figure illustrates three common types of brain tumors as observed in MRI scans: Glioma, Meningioma, and Pituitary tumor. These tumor types differ in origin, location, appearance, and clinical impact:

- ❖ **Glioma:** A tumor that originates from glial cells within the brain tissue. It often appears infiltrative and irregular in MRI images, making segmentation and classification challenging.
- ❖ **Meningioma:** A tumor that arises from the meninges (the protective layers surrounding the brain). It is usually well defined and located near the skull, often appearing as a clearly bounded mass in MRI scans.
- ❖ **Pituitary Tumor:** A tumor located in the pituitary gland at the base of the brain. It typically appears in a small, central region and may affect hormonal balance.

Medical Imaging:

Medical imaging is a field of medicine and technology that focuses on creating visual representations of the interior of the body for clinical analysis and medical intervention. It includes techniques such as MRI, CT scans, Xrays, and ultrasound. In artificial intelligence research, medical imaging provides critical data for developing automated diagnostic systems.

References

- [1] A. Hassan, M. Refaat and A. M. Hemeida, "Image classification based on Deep Learning," *Aswan University Journal of Sciences and*, pp. 11-35, 2022.
- [2] J. Wu, "A Review of Research on Image Classification Methods," *Proceedings of CONF-MLA Workshop: Mastering the Art of GANs*, pp. 110-115, 2024.
- [3] D. S. D. K, A. KV, K. S and J. S., "Data Augmentation using GANs for Image Classifications.," *Global Scientific Journals (GSJ)*, p. 522, 2025.
- [4] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, MIT Press, 2016.
- [5] T. K. Kurdis and A. M. A. Albrifkani, "Deep Learning Classification Algorithms Applications: A Review," *The Indonesian Journal of Computer Science (IJCS)*, no. 3, 2024.
- [6] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, Pearson, 2020.
- [7] C. Staff, "What Is Machine Learning? Definition, Types, and Examples," [Online]. [Accessed 28 04 2026].
- [8] Z. Y. H. D. M. D. Frank Emmert-Streib, "An Introductory Review of Deep Learning for Prediction Models With Big Data.," *Frontiers in Artificial Intelligence*, 2020.
- [9] A. Sherstinsky, "Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network," *Physica D: Nonlinear Phenomena*, 2020.
- [10] J. Schmidhuber, "Deep Learning in Neural Networks: An Overview," *Neural Networks*, 2015.
- [11] A. W. T. e. a. Creswell, "Generative Adversarial Networks: An Overview," *IEEE Signal Processing Magazine*, pp. 53-65, 2021.
- [12] Y. B. A. C. Ian Goodfellow, *Deep Learning*, 2016.
- [13] . Merry, J. Gyalmo, . Alom, . Aung and . A. Azim, "Transfer learning: a friendly introduction," *Journal of Big Data*, 2022.
- [14] H. E. C.-L. A. S. N. J. M. M. M. E. & G. T. Kim, "Transfer learning for medical image classification: A literature review," *BMC Medical Imaging*, 2022.
- [15] P. M. A. S. Deepak, "Brain tumor classification using deep CNN features via transfer learning," *Computers in Biology and Medicine*, 2019.

